

4. Научные статьи [Электронный ресурс] – Режим доступа: <https://nauchniestati.ru/bank/primery/nauchnaya-statya-na-temu-rol-elektronnyj-podpisi-v-rossii/> (дата обращения 23.06.2020).

5. Миронов И.Д. Правовое регулирование электронной подписи в России. - Иваново: Олимп, 2019. –С.50-55.

6. Проблемы долгосрочного хранения электронных деловых документов [Электронный ресурс] – Режим доступа: <https://www.top-personal.ru/officeworkissue.html?303> (дата обращения 24.06.2020).

7. Филонова О.И., Полякова Е.Н. Правовое обеспечение внедрения цифровых технологий в сфере электронных обращений граждан в органы государственной власти. - Тюмень: Вестник Тюменского государственного университета. Социально-экономические и правовые исследования, 2018. –С. 119-132.

**Замятин Константин Алексеевич**

ФГБОУ ВО «Курганский государственный университет»,  
студент кафедры «Безопасность информационных и  
автоматизированных систем»,  
reser09@mail.ru, Курган, Россия

**Москвин Владимир Викторович**

ФГБОУ ВО «Курганский государственный университет»,  
старший преподаватель кафедры  
«Безопасность информационных и автоматизированных систем»,  
bias@kgsu.ru, Курган, Россия

**Дик Дмитрий Иванович**

ФГБОУ ВО «Курганский государственный университет»,  
канд. техн. наук, доцент кафедры  
«Безопасность информационных и автоматизированных систем»,  
ddibox@mail.ru, Курган, Россия

## **ВОЗМОЖНОСТИ ТЕХНОЛОГИИ «БОЛЬШИЕ ДАННЫЕ» (BIG DATA)**

УДК 004.67

**Аннотация.** С 2014 года началось активное внедрение информационных технологий во многие отрасли страны. Для определения правильного вектора развития необходимо производить сбор статистических данных и их анализ. За большой период времени накапливается достаточно много статистических данных, которые тяжело обработать

стандартными средствами. Благодаря серии подходов и инструментов технологии Big Data, обработка таких данных облегчается.

**Ключевые слова:** Big Data, информация, данные, анализ.

**Annotation.** Since 2014, the active introduction of information technologies in many sectors of the country began. Statistics should be collected and analysed to determine the correct vector of development. Over a long period of time, quite a lot of statistics are accumulated, which are difficult to process by standard means. Thanks to a series of approaches and tools of Big Data technology, the processing of such data is facilitated.

**Keywords:** Big Data, information, data, analysis.

В ноябре 2013 года утверждена «Стратегия развития отрасли информационных технологий в Российской Федерации на 2014–2020 годы и на перспективу до 2025 года», в документе технология обработки Big Data обозначена в числе «прорывных для мировой индустрии» [1].

В связи со стремительным развитием информационных технологий и ростом числа гаджетов появляется необходимость в анализе полученной информации об изменении рынка, потребительском интересе. Если производить глобальный сбор данных, в результате накопится огромное количество информации, и будет расти по экспоненте. При этом источниками такой информации будут не люди, а вычислительные устройства. Это приборы для мониторинга, системы видеонаблюдения, персональные компьютеры, смартфоны, датчики, сенсоры и прочее [2].

Под термином Big Data скрывается большой объем информации, масштабы, разнообразие и сложность которых требует современную архитектуру, разработку новых алгоритмов и средств анализа для управления таким количеством информации. Главная задача Big Data – это обработка большого объема неструктурированных данных, систематизация этих данных, и их анализ, что приведет к выявлению закономерности в полученных данных, которую человек не в состоянии заметить.

Аналитика больших данных становится одной из наиболее востребованных задач в современном бизнесе. По оценкам компании Frost & Sullivan в 2021 году общий объем мирового рынка аналитики больших данных

увеличится по сравнению с показателями 2017 года более чем в 2 раза и составит \$64 млрд (рис. 1).



Рис. 1. Рост мирового рынка Big Data

При постоянном сборе информации может появляться «мусор», то есть данные, которые не пригодны для аналитики. Существует три характеристики [3], которые позволяют отнести информацию именно к Big Data:

- Volume (большой объем) – это большой объем информации, для обработки которой требуются значительные вычислительные ресурсы и новые подходы для анализа такой информации;
- Variety (разнообразие данных) – большая часть полученной информации не структурирована, не упорядочена и содержится в различных форматах;
- Velocity (скорость изменения данных) – данная характеристика указывает на растущую скорость накопления данных и на скорость обработки их, поэтому наиболее актуальны системы, которые производят анализ данных в реальном времени.

Однако, мало отнести полученную информацию к категории Big Data, с ней нужно правильно работать [3]. Существуют основные принципы для работы с такими данными:

- горизонтальная масштабируемость. Так как данные постоянно поступают в систему, может случиться так, что их попросту негде будет хранить, поэтому любая система, которая подразумевает обработку больших данных должно легко расширяться. Если выросло количество информации, то прямо пропорционально должно и вырасти количество железа в кластере;
- отказоустойчивость. Так как машин в кластере может быть большое количество, они по тем или иным причинам могут выходить из строя. Методы работы с большими данными должны учитывать подобные сбои, и продолжать работать без каких-либо значимых последствий;
- локальность данных. В больших распределенных системах информация может находиться на десятках и сотнях машин. В данный принцип заложено то, что информация должна обрабатываться на той же машине на которой она хранится, в противном случае затраты на передачу информации от одной машины к другой, могут превысить затраты на обработку этой самой информации. Технология Big Data активно используется во многих сферах деятельности. По результатам опроса «Используете ли вы технологии Big Data?», проведенного среди компаний, можно сделать вывод, что наиболее глубоко технология «Big Data» используется в отрасли телекоммуникации и инжиниринга (рис. 2).



Рис. 2. Внедрение технологии Big Data различных отраслях

В телекоммуникациях, а в частности у сотовых операторов, имеются весьма объемные базы данных, благодаря чему они имеют возможность производить глубокий анализ этой информации. Благодаря анализу данных они могут использовать полученные сведения для удержания клиентов, предложения лучших тарифов на основе их использования связи, или же привлечение новых на основе общих тенденции и конкурентной борьбы.

В 2014 году компания МТС начал тестировать технологию Big Data и тестировать ее для проверки нескольких бизнес гипотез. Через три года компания начала применять данную технологию во всех функциональных направлениях. Одних из основных источников данных – базовые станции, которые обслуживают 78 млн компонентов.

В финансовой сфере использование технологий Big Data позволяет проанализировать кредитоспособность клиента, ускорилось время рассмотрения кредитных заявок, раньше на это мог уходить день и более, то сейчас это возможно за 10 минут. Появляется возможность анализа использования банковских услуг и определения круга лиц, которым эти услуги более выгодны.

В России крупнейшей финансовой компанией используемой технологии Big Data, является компания «Сбербанк». Компания начал внедрять данную технологию с 2014 года, за 6 лет у них накопилось 120 ПБ данных. По заявлению официальных представителей, за секунду банк осуществляет порядка 12 тыс транзакций, и информация о них попадает в их базу данных. На основе этих данных более 200 команд разрабатывают банковские продукты и решения. Для работы с Big Data «Сбербанк» приобрел контрольный пакет в RuTarget. Компания так же использует программные продукты от Oracle и Teradata [4].

Одним из первопроходцев в использовании Big Data в России стала компания «Яндекс». Первая система процессинга Big Data в компании появилась в 2007 году. С ростом популярности сервисов «Яндекс» появилась

задача анализа использования этих сервисов. В 2006 году компания столкнулась с проблемой невозможности обработать большие объемы данных традиционными способами. Проблема заключалась в том, что не было технологии горизонтального расширения серверных систем. В настоящее время практически во всех продуктах «Яндекса» ведется сбор статистических данных для их анализа. В компании работают алгоритмы прогноза пробок, оптимизации поисковой выдачи, фильтрации спама, музыкальных рекомендации.

В 2014 году компания создала Yandex Data Factory. Это сервис, который специализируется на анализе больших данных и машинного обучения. Сервис предназначен для клиентов, которые хотят повысить эффективность своего предприятия за счет технологии искусственного интеллекта. Данным сервисом пользуются такие компании как, Intel, Магнитогорский металлургический комбинат, «Газпромнефть» и другие.

Однако, у технологии, которая активно внедряется во многих компаниях, есть и темная сторона, связанная с применением законодательства о персональных данных. Во-первых, законодательно закреплено, что обработка персональных данных должна осуществляться, исключительно с первоначально заявленными целями обработки и недопустимо объединение различных баз данных, данный факт вступает в противоречие с технологией Big Data и бизнес практикой. Во-вторых, применение технологии Big Data делает бесполезным обезличивание информации, и оно не может выполнять функцию средства защиты персональных данных [5].

### **Заключение**

Информация приобрела статус ценного актива и для правильного использования информации, нужно производить ее анализ. Анализ больших данных – ресурсоемкая задача, требующая для решения специальных инструментов и большую вычислительную мощность. Технологии Big Data позволяют собирать и анализировать данные в кратчайшие сроки и внедрение

ее в любые сферы бизнеса позволяют правильно распределять финансовые расходы.

### **Список использованной литературы**

1. Распоряжение Правительства РФ от 01.11.2013 N 2036-р «Об утверждении Стратегии развития отрасли информационных технологий в Российской Федерации на 2014–2020 годы и на перспективу до 2025 года». 33 с. URL: [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_154161/](http://www.consultant.ru/document/cons_doc_LAW_154161/) (Дата обращения 24.06.2020).
2. Соколов С. С., Новоселов Р. Ю., Митрофанова А. В. Методы обеспечения доступности информации в высоконагруженных информационных системах // Вестник УРФО. Безопасность в информационной сфере. 2018. № 2 (28). С. 31–35.
3. Веретенников А. В. BigData: анализ больших данных сегодня // Молодой ученый. 2017. № 32(166). С. 9–12.
4. Иванов П. Д., Вампилов В. Ж. Технологии Big Data и их применение на современном промышленном предприятии / П. Д. Иванов, В. Ж. Вампилов // Инженерный журнал: Наука и инновации. –2014. –№8. –С. 3.
5. Савельев А. И. Проблема применения законодательства о персональных данных в эпоху «Больших данных» (Big Data) // Право. Журнал Высшей школы экономики. 2015. № 1. С. 43–66.
6. Алексеева А. О., Сысолятин Р. М., Полякова Е. Н. Проблемы обеспечения безопасности больших данных // Наука и молодежь в XXI веке. Сборник научных трудов студентов и молодых ученых / ред. Роговая В.Г., Косовских С.В. –Курган, 2017. – С. 255–259.